



# Shortest-path constraints for 3D multiobject semiautomatic segmentation via clustering and Graph Cut

Razmig Kéchichian, Sébastien Valette, Michel Desvignes, Rémy Prost

## ► To cite this version:

Razmig Kéchichian, Sébastien Valette, Michel Desvignes, Rémy Prost. Shortest-path constraints for 3D multiobject semiautomatic segmentation via clustering and Graph Cut. IEEE Transactions on Image Processing, 2013, 22 (11), pp.4224-4236. 10.1109/TIP.2013.2271192 . hal-00983319

**HAL Id: hal-00983319**

**<https://hal.science/hal-00983319>**

Submitted on 25 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Shortest-Path Constraints for 3D Multiobject Semiautomatic Segmentation via Clustering and Graph Cut

Razmig Kéchichian, Sébastien Valette\*, Michel Desvignes, and Rémy Prost, *Member, IEEE*

**Abstract**—We derive shortest-path constraints from graph models of structure adjacency relations and introduce them in a joint centroidal Voronoi image clustering and Graph Cut multiobject semiautomatic segmentation framework. The vicinity prior model thus defined is a piecewise-constant model incurring multiple levels of penalization capturing the spatial configuration of structures in multiobject segmentation. Qualitative and quantitative analysis and comparison with a Potts prior-based approach and our previous contribution on synthetic, simulated and real medical images show that the vicinity prior allows for the correct segmentation of distinct structures having identical intensity profiles and improves the precision of segmentation boundary placement while being fairly robust to clustering resolution. The clustering approach we take to simplify images prior to segmentation strikes a good balance between boundary adaptivity and cluster compactness criteria furthermore allowing to control the trade-off. Compared to a direct application of segmentation on voxels, the clustering step improves the overall runtime and memory footprint of the segmentation process up to an order of magnitude without compromising the quality of the result.

**Index Terms**—Image segmentation, image clustering, Markov random field, spatial prior, Graph Cut.

**EDICS categories**—TEC-BIP, SMR-SMD, SMR-STM.

## I. INTRODUCTION

COMPUTER-AIDED medical image analysis has traditionally focused on single-organ or pathology applications. Recent technological and algorithmic advances have brought increasing interest in simultaneous analysis and segmentation of multiple anatomical structures for comprehensive diagnosis and preoperative planning [1]–[4]. Moreover, the creation of full-body patient-specific models for the semantic navigation of anatomy is becoming a popular application [5].

Clinical practice, especially whole-body Computed Tomography (CT) scanning, often generates large numbers of high-resolution images, which makes tasks of efficient data access, transfer, analysis and visualization challenging, especially in

today's distributed computing environments which have seen increasing use of handheld terminals for interactive data access and visualization of anatomy.

CT and Magnetic Resonance (MR) images have intrinsic characteristics that render automatic analysis particularly difficult. In theory, they are regarded as piecewise-constant intensity maps over a number of tissue classes. However, contrast between tissue class intensities, which facilitates segmentation, depends on imaging conditions which are seldomly ideal. In practice, the piecewise-constancy property is considerably degraded by various noise sources and the partial volume effect (multiple tissue-class occupancy within a voxel) which creates diffuse edges making it difficult to identify true structure boundaries. MR images may, in addition, suffer from spatial distortion of tissue intensity due to main magnetic field inhomogeneity. Regardless of the imaging modality and related artifacts, many anatomically and functionally distinct structures can have similar intensity levels in images and, furthermore, blend into surrounding tissues having intensities close to their own. It is impossible to identify and segment such structures on the basis of intensity information only.

Semiautomatic segmentation methods are usually resorted to in order to alleviate prohibitive time and effort requirements of manual delineation and as an alternative to fully-automatic methods when the latter are not applicable, for example due to the singularity of the studied case or the lack of training sets. Furthermore, automatic methods which rely on image and anatomical features learned beforehand from presegmented data, while effective in many cases, are prone to be bound to the characteristics of the training set and its variabilities, and may therefore have difficulty in accounting for previously unencountered but possible pathological cases and photometric characteristics of images acquired under different conditions.

A wide variety of semiautomatic segmentation approaches relying on different interaction paradigms have been proposed (see references in [6], [7]) where a reasonable amount of user input can initialize and steer the segmentation process refining its result by resolving erroneous decisions made by the algorithm. In particular, when segmentation is formulated as a voxel labeling problem where each segmented object is represented by a label class, user interaction can be carried out by the attribution of “seed” labels to voxels inside targeted structures. This provides clues on what the user intends to segment and can be used, for example, to collect intensity

Manuscript received on September 26, 2012; revised on January 21 and April 10, 2013. This work is supported by the Région Rhône-Alpes (France) via the SIMED project of the research cluster ISLE.

R. Kéchichian, S. Valette and R. Prost are with Creatis research center: Université de Lyon; CNRS UMR 5220; Inserm U1044; INSA-Lyon; Université Lyon 1, Villeurbanne, France (e-mail: firstname.lastname@creatis.insa-lyon.fr).

M. Desvignes is with Gipsa-lab: CNRS UMR 5216; Grenoble-INP; Université Joseph Fourier; Université Stendhal, Saint-Martin-d'Hères, France (e-mail: michel.desvignes@gipsa-lab.grenoble-inp.fr).

Digital Object Identifier

statistics of targeted objects and their relative positions, and to constrain the solution space of the algorithm by prohibiting the attribution of other labels to seed-marked voxels.

#### A. Graph Cut in segmentation and its complexity

A fundamental problem of image processing and computer vision, image segmentation is the back-bone of many applications, especially in medical imaging. Graph Cut methods, which have been widely applied to single-object segmentation problems [8], rely on a maximum-flow binary global optimization scheme capable of finding the optimal solution in mathematical sense [9]. With few exceptions [10], multiobject generalizations of this approach do not, in general, have such optimality properties due to the NP-hardness of the underlying problem, but some algorithms [11] are nevertheless capable of finding a local solution optimal within a known factor of the global optimum. Simultaneous multiobject segmentation approaches, on the other hand, are superior to their sequential counterparts in that they raise questions neither on the best segmentation sequence to follow nor on how to avoid the propagation of errors on individual segmentations [2].

Graph Cut algorithms usually employ rectangular grid graphs for data structure. In applications on high-resolution 3D medical images where several tens of millions of voxels are common, such graphs have to be defined on a number of vertices equal to the number of image voxels and a number of edges at least five times as much. With data structure size touching on the limit or exceeding the storage capacity of computer memory, numerical optimization becomes impractical.

Approaches which address the complexity of Graph Cut methods on high-resolution data can be broadly classified into two categories; algorithm parallelization and image simplification. Parallel versions of some maximum flow algorithms have been devised for multiprocessor architectures [12] and graphics processing units [13]. While these methods attain good speedups, they do not reduce memory footprint since they operate on the entire image. Image simplification approaches address this problem directly observing that the pixel representation is often redundant because objects usually comprise many similar pixels that could be grouped. In [14], Graph Cut segmentation is applied to a low-resolution image propagating the solution to the full resolution where Graph Cut segmentation is applied in a narrow band surrounding the projected object boundary. In order to accelerate user feedback in interactive segmentation, [15] oversegment the image via the Watershed algorithm. Graph Cut segmentation is then formulated according to the watershed regions graph. Starting from a grid partition, [16] cluster image pixels by an iterative k-means algorithm augmented by color similarity and shape compactness criteria. The cluster graph is then partitioned based on color information yielding a coarse segmentation of the image. Recent conditional random field-based (CRF) segmentation approaches such as [17], [18] rely on image clustering not only to reduce memory overhead but also to collect image features from clusters and their neighborhood.

#### B. Prior information in Graph Cut segmentation

Most successful segmentation methods usually incorporate some image or domain-specific prior information. Medical image segmentation methods should ideally be able to incorporate a wide range of anatomical and physiological prior information to ensure the consistency of the result with respect to anatomical properties and variabilities thereof. From the standpoint of Graph Cut optimization, prior information can be regarded as a set of constraints defined by a reference model.

Many forms of prior information have been introduced in Graph Cut segmentation for different applications. We broadly classify the most relevant of such approaches in three categories according to the nature of the introduced prior.

*a) Location prior*, where spatially-varying prior probability maps of structure locations, also called probabilistic atlases, are used to guide segmentation [4], [19]. In simplest cases, these are defined as distance maps [20]. Such probabilistic maps need to be learned in advance and most have to be registered to an image before it is segmented. Registration cannot trivially handle topological changes, due to pathologies for example, and is often computationally intensive.

*b) Shape prior*, where explicit or implicit shape constraints are used to match the segmented object with predefined or learned shapes, such as in [21] where an affine-invariant shape similarity measure is used in sequential multiobject Graph Cut segmentation, and in [22] where a statistical model for a shape and its variabilities is learned from a set of training shapes via kernel principal component analysis. Shape-prior segmentation approaches entail shape location initialization and alignment mechanisms, and often require model learning.

*c) Spatial configuration prior*, the use of which arises in the context of multiobject segmentation, where (usually pairwise) relations define interobject geometric and spatial relationships. It has appeared in scene layout estimation and segmentation where spatial relationships are predefined on a limited number of classes [23], and in part-based object detection and segmentation where such relationships are learned from annotated training data [18], [24]. [25] segment multiple or compound objects encountered in relative positions of containment and exclusion, also introducing constraints on preferred distances between the boundaries of regions. Some useful configurations however cannot be represented by definitions therein, such as the inclusion of an adjacent pair of objects in a third.

There are a number of prior introduction approaches closely related to our work which have appeared outside Graph Cut segmentation literature. [26] use a simple three-level piecewise-constant prior model for tissue classification in brain MR images. According to a brain tissue model, it encourages the identical classification of pairs of neighboring pixels, allowing different class adjacency patterns which are consistent with the model while penalizing adjacencies which are not. We use a similar prior for Graph Cut multiobject segmentation in our earlier work [27]. In [1], pairwise geometric relationships are defined by an anatomical network on easily detectable body landmarks and used in the organ detection phase of a sequential segmentation framework for thoracic and abdominal organs in whole-body CT scans. [3] address the same problem

with an automatic two-stage method whereby once relatively stable organs are identified and segmented, organ interrelations are used to proceed with the segmentation of variable organs by fitting previously learned statistical organ atlases and shape models. Organ interrelations are defined according to canonical correlation analysis and represented by a directed graph. Intuitively, these relations represent the degree of influence of an organ's surface upon another. A generic, graph-based scene model representing spatial information explicitly is also used by [2] in the problem of segmenting multiple subcortical structures in brain MR images. Fuzzy representations of distance, adjacency and spatial directionality relations are used to account for model imprecisions and variabilities. However, the segmentation algorithm is sequential and requires error detecting and segmentation order rearranging strategies.

### C. Contribution

We propose a piecewise-constant prior model incurring multiple levels of penalization capturing the spatial configuration of structures in multiobject segmentation. The vicinity prior, as we term it, is defined as shortest-path pairwise constraints on a graph model of interobject adjacency relations.

We introduce the vicinity prior, along with user-supplied constraints, into a joint centroidal Voronoi image clustering and Graph Cut multiobject semiautomatic segmentation framework. The clustering method itself we use to simplify images prior to segmentation provides a good balance between the conflicting goals of boundary adaptivity and cluster compactness furthermore allowing to control the trade-off. Compared to a direct application of segmentation on voxels, the clustering step improves overall runtime and memory footprint of the segmentation process up to an order of magnitude without compromising the quality of segmentation in practice.

Qualitative and quantitative analysis and comparison with a standard Potts prior-based approach and our previous method [27] on synthetic, simulated and real medical images confirm the advantages of the vicinity prior particularly in the correct segmentation of distinct structures having similar intensities, the accurate placement of segmented structure boundaries even with coarsely clustered images subject to high levels of noise, and the robustness of segmentation to clustering resolution.

### D. Paper organization

Section II outlines the image-adaptive centroidal Voronoi tessellation and compares it with two closely-related algorithms [28], [29]. The statistical segmentation approach is described in Section III. In Section III-A, we present our graph-based vicinity prior model followed by data-cost definition in Section III-B. We briefly describe the numerical optimization method we employ in Section III-C and establish the necessary condition of its applicability. Results of qualitative and quantitative evaluation on synthetic, simulated and real medical images are given in Section IV. In Section IV-A we present evaluation results on a dataset generated from our synthetic phantom (Fig. 3a). Next, in Section IV-B1 we give quantitative evaluation results on the BrainWeb simulated MRI dataset [30] and compare them to brain tissue classification results reported

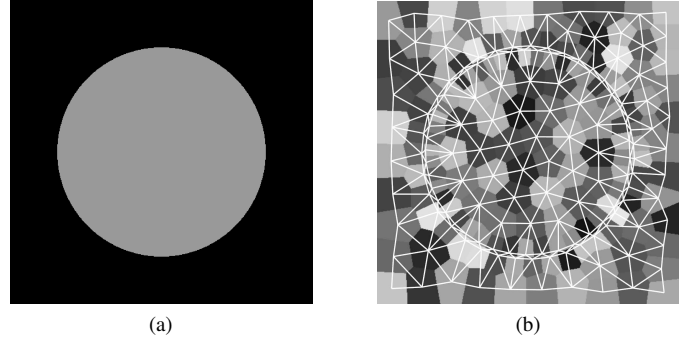


Fig. 1. A CVT clustering and its dual graph for a circle image.

in [31]. Finally, in Section IV-B2 we present a segmentation case study comprising 34 thoracic and abdominal structures in a CT image and conclude the paper in Section V with a brief discussion outlining future venues of research.

## II. IMAGE CLUSTERING BY CENTROIDAL VORONOI TESSELLATION

Several image clustering algorithms have been developed by vision and pattern analysis communities for image classification and segmentation. Of special interest is a class of algorithms which produce a dense oversegmentation of compact clusters, often called superpixels, having relatively uniform size and shape, which furthermore adapt to local intensity edges. Two recent examples are the TurboPixels algorithm of [28] and the Graph Cut superpixels algorithm of [29] which achieve a better balance between the conflicting goals of compactness and boundary adherence than some well-known image partitioning algorithms which produce segments of highly variable shape and size, like Watershed [32] and Mean-shift [33] algorithms. A large segment of irregular shape is more likely to span more than one object, especially in the absence of boundary cues with insufficient contrast. On the other hand, a compact regular shape is less likely to cross object boundaries unless they are too wiggly. However, compactness comes at the expense of boundary adherence, it is therefore desirable to be able to control the trade-off.

Before presenting our clustering approach, let us define a grayscale image  $\mathcal{I}$  as a mapping  $\mathcal{I}: V \mapsto I$  from a voxel domain  $V = \{v \mid v = (x, y, z)\}$  to a set of gray levels  $I \subset \mathbb{R}$ . To simplify notation, we shall write  $v \in \mathcal{I}$ , and denote the gray-level of a voxel by  $I_v$ .

Given a grayscale image  $\mathcal{I}$  and  $n$  sites  $c_i \in \mathcal{I}$  such that  $1 \leq i \leq n$ , a Voronoi tessellation partitions the image  $\mathcal{I}$  into  $n$  disjoint clusters  $C_i$  associated with each site  $c_i$ :

$$C_i = \{v \in \mathcal{I} \mid d(v, c_i) < d(v, c_j); 1 \leq j \leq n, j \neq i\}. \quad (1)$$

In a centroidal Voronoi tessellation (CVT), each site  $c_i$  corresponds to the mass centroid of the associated cluster, and the tessellation minimizes the following energy:

$$E(v; c_i) = \sum_{i=1}^n \left( \sum_{v \in C_i} \rho(v) \|v - c_i\|^2 \right) \quad (2)$$

where  $\rho(v)$  is a density function defined below.

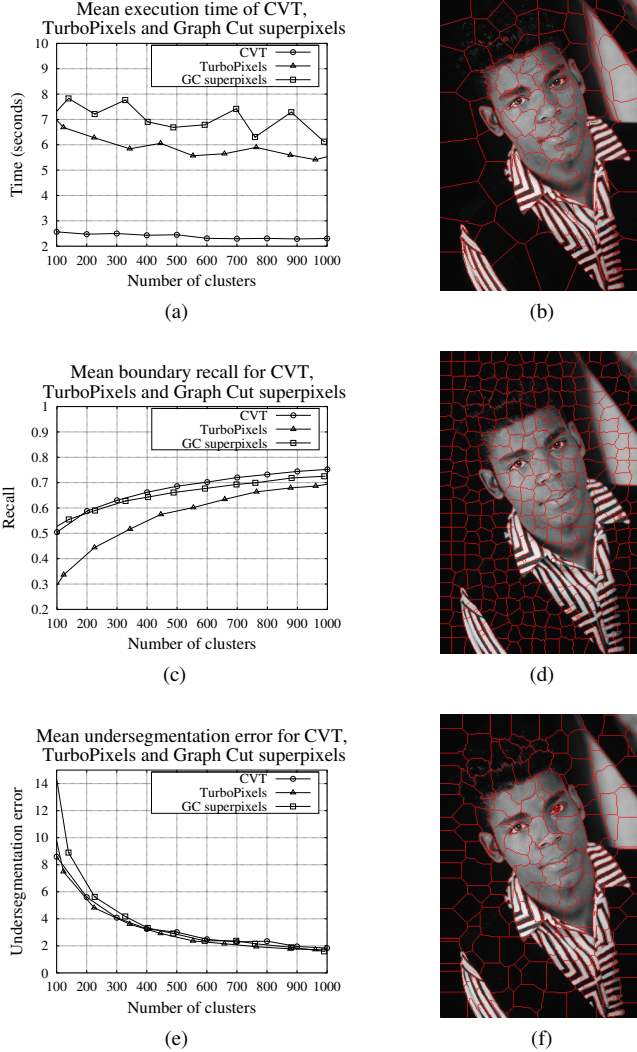


Fig. 2. Left column: quantitative evaluation of clustering algorithms on the Berkeley segmentation dataset. Right column: qualitative comparison of their output with 300 clusters on a single photo from the dataset. Boundary recall and undersegmentation error are respectively (b) CVT: 0.59, 1.53 (d) TurboPixels: 0.52, 2.01 (f) GC superpixels: 0.59, 2.76.

We extend the cluster geometry compactness property expressed by (2) by adding an intensity-space norm as well:

$$F(v; c_i) = \sum_{i=1}^n \left( \sum_{v \in C_i} \rho(v) (\|v - c_i\|^2 + \alpha \|I_v - I_i\|^2) \right) \quad (3)$$

where  $\alpha$  is a positive scalar and  $I_i$  is the gray-level of the cluster  $C_i$  defined as the mean intensity of its voxels. Intuitively, minimizing (3) now corresponds to maximizing cluster compactness in terms of both geometry and gray-level.

In (3), we define  $\rho(v)$  as a linear function of intensity-gradient magnitude at voxel  $v$ ,  $\rho(v) = m|\nabla I_v| + b$ . This encourages the alignment of clusters with intensity edges in the image and the formation of relatively small clusters nearby, thus allowing fine-grained placement of segmentation boundaries. In all results reported in this paper, we set  $\alpha = m = b = 1$ . Refer to Fig. 1b for an illustration in 2D.

To minimize (3), we use a variant of the clustering algorithm

in [34] which approximates a CVT in a computationally-efficient manner, involving only local queries on voxels located on boundaries of pairs of clusters. To assess its runtime performance and output quality, we have compared it with the TurboPixels and Graph Cut superpixels algorithms, both of which have similar cluster compactness criteria, on the Berkeley segmentation dataset [35]. This dataset comprises 300 generic photos of size  $321 \times 481$  segmented by different human subjects. In quantitative analysis, we evaluate runtime, boundary recall and undersegmentation error on 80 grayscale images from the dataset, choosing, for each such image, one ground-truth segmentation that does not contain small segments of few pixels. Such segments can produce large outliers and disrupt mean undersegmentation error measurements.

Boundary recall is defined as the fraction of ground-truth boundary pixels which fall within a small distance from at least one cluster boundary. We use a distance of one pixel since we seek high precision in segmentation boundary placement in medical applications. Given a CVT clustering of an image  $\mathcal{C} = \{C_1, \dots, C_n\}$  and a ground-truth partition of it  $\mathcal{T} = \{g_1, \dots, g_k\}$ , the undersegmentation error for segment  $g_j$  is:

$$\text{segerr}_j(\mathcal{C}, \mathcal{T}) = \frac{\left( \sum_{i: C_i \cap g_j \neq \emptyset} |C_i| \right) - |g_j|}{|g_j|} \quad (4)$$

For a given clustering, we average (4) over all segments of the corresponding ground-truth partition and give the mean error.

Figs. 2a, 2c and 2e present results of quantitative evaluation of the CVT clustering algorithm, the TurboPixels algorithm and the Graph Cut superpixels algorithm variant using cluster intensity homogeneity constraints. For a given clustering resolution, we report mean performance measures over all images. We can see from the graphs that the CVT clustering algorithm outperforms both TurboPixels and Graph Cut superpixels algorithms in terms of runtime and produces output of comparable quality to the best performing algorithm of the remaining two. Qualitative comparison of the output of the three clustering algorithms on a Berkeley dataset photo is given in Figs. 2b, 2d and 2f. We note that evaluations were conducted on a 4-core processor running at 2.84 GHz using software implementations provided by respective authors of the three algorithms.

For discussion purposes in subsequent sections, we define the dual graph of a CVT, illustrated in Fig. 1b. In 2D, denote the boundary (the surface, in 3D) of a cluster  $C_i$  by  $\partial C_i$ . Given a CVT clustering  $\mathcal{C}$  where clusters are indexed by  $\mathcal{S}$ , let  $\mathcal{G} = \langle \mathcal{S}, \mathcal{E} \rangle$  be an undirected graph on cluster centroids where pairs of clusters having nonzero length (area) common boundary (surface) define the set of edges  $\mathcal{E} = \{\{i, j\} \mid i, j \in \mathcal{S}, |\partial C_i \cap \partial C_j| \neq 0\}$ . Consequently, the neighborhood of a node  $i \in \mathcal{S}$  is defined as  $\mathcal{N}_i = \{j \mid j \in \mathcal{S}, \exists \{i, j\} \in \mathcal{E}\}$ .

### III. MULTIOBJECT SEGMENTATION AS BAYESIAN LABELING

We shall formulate segmentation as a labeling problem, defined as the assignment of a label from a set of  $m$  labels  $L = \{l_1, \dots, l_m\}$  representing objects to be segmented to each of the variables in a set of  $n$  variables corresponding to the clusters of a CVT-clustered image  $\mathcal{C} = \{C_1, \dots, C_n\}$

indexed by  $\mathcal{S}$ . Assume furthermore that each variable  $i \in \mathcal{S}$  is associated with the corresponding node in the dual graph  $\mathcal{G}$  of the CVT defined at the end of Section II.

An assignment of labels to all variables is called a configuration, and is denoted by  $\ell \in \mathcal{L}$ , where  $\mathcal{L} = L^n$  is the space of all possible configurations. The assignment of a label to a single variable  $i \in \mathcal{S}$  is denoted by  $\ell_i$ . The space of admissible configurations, which corresponds to possible segmentations, can be identical to  $\mathcal{L}$  or, if user-specified constraints are taken into account, to a subset of  $\mathcal{L}$ . Let  $X = \{X_1, \dots, X_n\}$  be a set of random variables on  $\mathcal{S}$  taking its values in  $L$ . A configuration  $\ell$  can also be expressed in probabilistic terms as the joint event  $X_1 = \ell_1, \dots, X_n = \ell_n$ , abbreviated as  $X = \ell$ .

The “best” segmentation is defined as the configuration corresponding to the highest mode of the posterior probability distribution of configurations  $\Pr(X = \ell \mid O)$  conditional on some image-derived observation  $O$ . According to the Bayes rule, the posterior probability can be stated in terms of a likelihood distribution  $\Pr(O \mid X = \ell)$  and a prior distribution of configurations  $\Pr(X = \ell)$ . If the latter can be expressed as a first-order Markov random field (MRF) for label configurations with respect to  $\mathcal{G}$  and image-derived likelihood densities can be defined for each label  $l \in L$  from user-supplied evidence, then the maximum probability configuration can be equivalently found by minimizing the following energy function [36]:

$$E(\ell) = t \sum_{i \in \mathcal{S}} D_i(\ell_i) + \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{N}_i} V_{i,j}(\ell_i, \ell_j) \quad (5)$$

where  $t$  is a hyperparameter called the temperature and  $\mathcal{N}_i$  is the neighborhood of the variable  $i \in \mathcal{S}$ . Pairwise terms of (5) encode prior information on interactions between labels assigned to pairs of neighboring sites encouraging the spatial consistency of labeling with respect to a reference model. Unary terms, also called data terms, are negative log-likelihood functions derived from observed data and measure the cost of assigning a label to respective variables.

#### A. The vicinity prior

For image segmentation, the pairwise interaction prior in (5) has to encourage piecewise-constant labeling, partitioning the image into several regions where voxels within a region are identically labeled. Furthermore, it has to be discontinuity preserving, in that it should attribute identical labels only to image regions which do not contain intensity edges.

The earliest discontinuity-preserving piecewise-constant prior model is the Ising model in binary labeling where  $|L| = 2$  [37]. It is defined as  $V_{i,j}(\ell_i, \ell_j) = 1 - \delta(\ell_i, \ell_j)$ , where  $\delta$  is the Kronecker delta. Its generalization for multiple labels where  $|L| > 2$  is called the Potts model, which has widely been used in image segmentation and continues to enjoy popularity in recent CRF-based approaches [17], [38]. We can see that this model incurs a single level of penalization corresponding to the attribution of different labels to pairs of neighboring variables and incurs no penalty for the identical labeling of such pairs. This penalization scheme stems from the very nature of binary labeling where only a single type of boundary exists between regions labeled by either label.

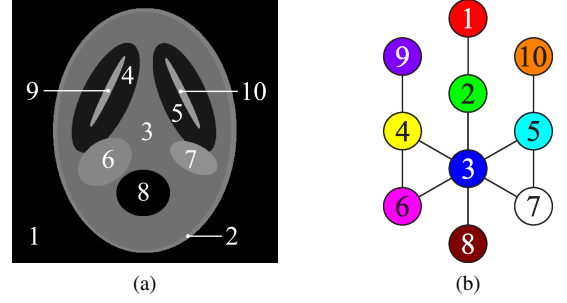


Fig. 3. (a) Synthetic phantom and (b) the graph model for adjacency relations of its structures.

We extend the definition of the Potts prior model to account for multiple types of boundaries in multioobject segmentation, incurring multiple levels of penalization. To motivate our approach, let us first give an illustration on a synthetic example. Consider the image in Fig. 3a. It is further described in Section IV-A where we give evaluation results that use it as ground-truth estimate. Notice how the structure labeled “9” is nested in structure “4” which, in turn, is adjacent to structure “6”. Due to this adjacency pattern, it makes perfect sense to encourage the assignment of the pair of labels  $\{9, 4\}$  to neighbor variables in a segmentation task while penalizing the attribution of the pair  $\{9, 6\}$ . Hence,  $\forall i, j \in \mathcal{S}, V_{i,j}(9, 6) > V_{i,j}(9, 4)$ . Furthermore, since the structure “8” is farther still, we can say  $V_{i,j}(9, 8) > V_{i,j}(9, 6) > V_{i,j}(9, 4)$ . By contrast, the simple structure adjacency prior models employed in our earlier work [27] and in [26] penalize nonadjacent pairings  $\{9, 6\}$  and  $\{9, 8\}$  equally implying that they are equally unfit.

Recall that the set of labels  $L$  represents the objects to be segmented. Let  $\mathcal{R}$  be the set of symmetric binary relations on pairs of distinct labels,  $\mathcal{R} = \{r \mid l_a r l_b, l_a, l_b \in L, a \neq b\}$ , representing such relations as adjacency, distance, area of common surface, etc. Assume that each relation  $r \in \mathcal{R}$  has a corresponding characteristic function  $\hat{r}$  which gives a quantitative measure for  $r$ ;  $\hat{r}: L \times L \setminus \{(l_a, l_a)\} \mapsto \mathbb{R}^+$ .  $\mathcal{R}$  can be equivalently represented by a weighted undirected graph  $\mathcal{A} = \langle L, W \rangle$  on labels  $L$  with the set of edges  $W = \{\{l_a, l_b\} \mid \exists r \in \mathcal{R}, l_a r l_b, a \neq b\}$  where edge weights are defined by  $w(\{l_a, l_b\}) = \hat{r}(l_a, l_b)$ , such that  $w(\{l_a, l_b\}) = \infty, \forall r \in \mathcal{R}, l_a r l_b$ .

Given the graph  $\mathcal{A}$ , we define the pairwise term in (5) as

$$V_{i,j}(\ell_i, \ell_j) = \varpi(l_a, l_b), \quad \ell_i = l_a, \ell_j = l_b. \quad (6)$$

where  $\varpi(l_a, l_b)$  is the shortest-path weight from  $l_a$  to  $l_b$  in  $\mathcal{A}$ .

At present, we let relations  $r \in \mathcal{R}$  represent structure adjacency. A structure is considered to be adjacent to another in 2D if they share a boundary (a surface, in 3D). For two such structures  $a$  and  $b$ , the obvious definition of the adjacency relation’s characteristic function is  $\hat{r} = 1$  when  $\exists r \in \mathcal{R}, l_a r l_b$ . This results in a connected undirected graph  $\mathcal{A}$  with unit-weight edges between adjacent vertices. Shortest paths from a vertex to all others, conveniently calculated by an application of breadth-first search to the graph  $\mathcal{A}$ , define the *vicinity* of the associated structure with respect to all other structures, which is introduced in the energy function of segmentation (5) via pairwise interaction terms (6). Clearly, the graph  $\mathcal{A}$  constitutes



a prior model of explicit spatial configuration of objects. As an example, the graph prior model for the image structures in Fig. 3a is given in Fig. 3b.

Note that the graph  $\mathcal{A}$  is not limited to unit-weight edges. It can encode any other quantity that accumulates linearly along a path that we would want to minimize.

In 2D, we weigh (6) by the length (area, in 3D) of the common boundary (surface) of adjacent clusters  $|\partial C_i \cap \partial C_j|$  so that the sum of pairwise energies in (5) for any pair of labels  $l_a, l_b \in L$  is equal to the length (area) of the common boundary (surface) between the corresponding pair of objects multiplied by  $V(l_a, l_b)$ . Along with our data energy definition (7), this ensures that the energies of identical label configurations on CVT clusterings of different resolution are equal. Due to limited space, we cannot give the formal proof. Nevertheless, we establish the advantage of this weighting scheme in Section IV-A by quantitative comparison with the inverse distance weighting method frequently encountered in MRF-based segmentation, for example in [8] and [26].

### B. Likelihood-based data cost

The likelihood probability for a variable  $i \in \mathcal{S}$  given the label assignment  $\ell_i$  is given by  $\Pr(O_i = I_i | X_i = \ell_i)$ , where the observation  $O$  constitutes the spatially distinct records of image intensity levels for each variable. Given user-supplied “seeds”,<sup>1</sup> we estimate the conditional probability distribution of intensity  $\Pr(I | L = l)$  for every label  $l \in L$  as a Gauss-smoothed and normalized intensity histogram.

To facilitate the understanding of different cases, we define the data term  $D_i(\ell_i)$  of the energy (5) as follows:

$$D_i(\ell_i) = \begin{cases} 0, & \exists s_k \in C_i, k = \ell_i & (7a) \\ \infty, & \exists s_k \in C_i, k \neq \ell_i & (7b) \\ -\ln \prod_{v \in C_i} \Pr(I_v | \ell_i), & \nexists s_k \in C_i, \forall k \in L & (7c) \end{cases}$$

where  $s_k$  represents a voxel marked with a seed of label type  $k \in L$ . Note that (7) is a continuous function where (7a) and (7b) correspond to negative log-likelihood values at maximum and zero probability respectively. We represent the infinite cost in (7b) by a large constant several orders of magnitude ( $10^9$ – $10^{12}$ ) greater than the value of (7c) corresponding to the smallest estimated likelihood probability, so that, for practical settings of  $t > 0$ , (7a) and (7b) amount to imposing hard constraints for acceptable configurations. By comparison, [8] avoid multiplying (7b) by  $t$  in (5) and define a constant representing its cost such that prior terms do not accidentally impose a labeling decision on corresponding variables. In our formulation, as  $t$  greatly decreases, (7b) approaches the order of magnitude of prior terms until the data sum of (5) vanishes at  $t = 0$ , which is hardly an interesting setting in practice.

In subsequent refinements of an initial segmentation, we allow the user to introduce additional “corrective” seeds to constrain the label preference of incorrectly segmented image regions. In order not to disrupt initial appearance statistics, we do not take such seeds into account for likelihood estimation.

<sup>1</sup>Recall that seeds are label attributions to voxels inside targeted structures.

### C. Multilabel optimization via Graph Cut

Having defined the prior model and data-cost terms, we now turn our attention to the numerical optimization of the energy function of segmentation (5). We do not give the details of the optimization algorithm itself, but establish a necessary and sufficient condition of its applicability for optimizing the aforementioned energy function using our prior definition (6).

In [11], it has been shown that the multilabel minimization of (5) is NP-hard even with the simple discontinuity-preserving Potts prior model. In binary labeling with  $|L| = 2$  using this model, the global minimum of (5) can be computed in polynomial time by a maximum flow algorithm [9]. In general, however, and in our multiobject segmentation problem with  $|L| > 2$ , efficient approximation algorithms must be resorted to. We use the Expansion Moves multilabel Graph Cut algorithm described in [11], which has been shown to outperform popular multilabel optimization algorithms on benchmarks in terms of both speed and quality of obtained solutions [39].

The Expansion Moves algorithm produces a labeling locally-optimal within a known factor of the global minimum equal to  $2 \times \max V_{i,j}(l_a, l_b) / \min V_{i,j}(l_a, l_b)$ ,  $\forall i, j \in \mathcal{S}, l_a, l_b \in L, a \neq b$ . For example, this factor is equal to two for the Potts model. According to [40], the algorithm can minimize (5) if the pairwise term is submodular, that is, if the following condition holds  $\forall i, j \in \mathcal{S}, \forall l_a, l_b, l_c \in L$ :

$$V_{i,j}(l_a, l_a) + V_{i,j}(l_b, l_c) \leq V_{i,j}(l_a, l_c) + V_{i,j}(l_b, l_a). \quad (8)$$

The vicinity prior (6) is submodular by definition. This can be easily verified by observing that the shortest-path distance between any three vertices on the graph  $\mathcal{A}$  satisfies the triangle inequality. By comparison, the adjacency prior in [27] is essentially limited to definitions where the penalty of nonadjacency has to be twice the penalty of adjacency so that the prior is submodular. In other words, it is impossible to penalize nonadjacencies by some large constant without penalizing adjacencies half as severely.

## IV. EVALUATION RESULTS AND APPLICATIONS

This section presents the results of qualitative validation and quantitative analysis of our vicinity prior-based Graph Cut segmentation approach. In all evaluations, comparisons are drawn with the Potts prior-based approach, and in some cases with our earlier adjacency prior [27]. We note that our software implementation uses the multilabel energy optimization library developed by the authors of [11], [40], [41].

In all evaluations, we measure the quality of segmentation with respect to ground-truth via the well-known Dice similarity metric (DSM), which measures the overlap between the segmentation and the ground-truth. Let  $S_l$  and  $T_l$  represent the sets of voxels labeled with  $l \in L$  in the segmented  $\mathcal{I}$  and the ground-truth  $\mathcal{T}$  images respectively, and denote the DSM for label class  $l \in L$  by  $\text{dsm}_l(\mathcal{I}, \mathcal{T})$ . We define an “overall” segmentation quality metric, calculated from mean weighted DSM measures for all labels in  $L$  as:

$$\text{dsm}_L(\mathcal{I}, \mathcal{T}) = \frac{\sum_{l \in L} \text{dsm}_l(\mathcal{I}, \mathcal{T}) |T_l|}{|\mathcal{T}|}. \quad (9)$$

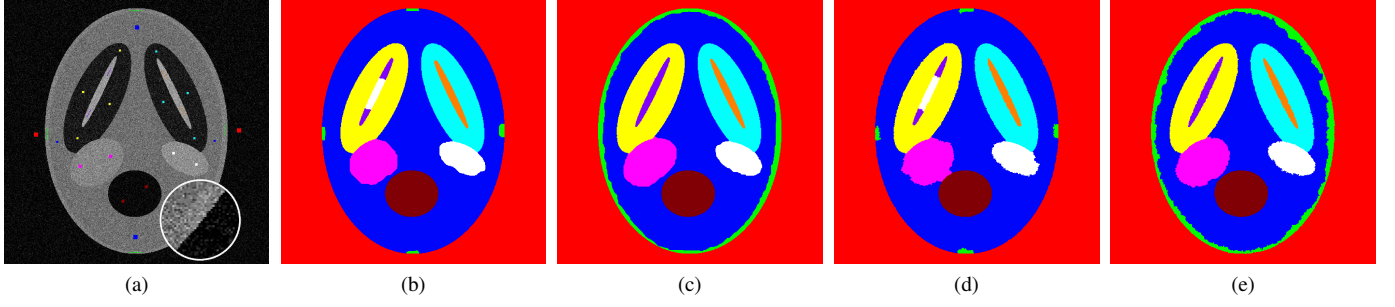


Fig. 4. (a) Image generated from the phantom in Fig. 3a with additive Gaussian noise  $\sigma = 0.09$  along with “seeds” used in all evaluations (b) Potts prior (DSM = 0.937) and (c) vicinity prior (DSM = 0.994) segmentations on pixels of the image (d) Potts prior (DSM = 0.938) and (e) vicinity prior (DSM = 0.986) segmentations on a 10% clustering of the image. Results correspond to best temperature settings. Refer to color code in Fig. 3b

In order to understand how reliable the segmentation algorithm is with either prior in attaining a given overall DSM value  $d$ , we evaluate a reliability function of DSM over the interval  $[0, 1]$  on a set of segmented images  $\mathbb{I}$ . Define the subset of images attaining an overall DSM level superior to  $d$  as  $\mathbb{I}_d = \{\mathcal{I} \in \mathbb{I} \mid \text{dsm}_L(\mathcal{I}, \mathcal{T}) > d\}$ , and let  $D$  be a random variable on  $[0, 1]$ . The reliability at DSM value  $d$  is given by:

$$\text{rel}(d) = \Pr(D > d) = \frac{|\mathbb{I}_d|}{|\mathbb{I}|}. \quad (10)$$

For the purpose of evaluating the precision of segmentation boundary placement on the medical case study in Section IV-B2, we use an error metric to measure the Euclidean distance of each point on the surface of a segmented structure from the surface of the corresponding ground-truth structure. Let  $M_S^l$  and  $M_T^l$  be respectively the triangle meshes representing isosurfaces of structure volumes labeled by  $l \in L$  in the segmented  $\mathcal{I}$  and the ground-truth  $\mathcal{T}$  image. We obtain such meshes by applying the Marching Cubes algorithm to labeled volumes [42]. The error at  $x \in M_S^l$  is defined as:

$$\text{err}(x) = \inf \{d(x, y) \mid y \in M_T^l\}. \quad (11)$$

#### A. Synthetic images

We have generated a dataset from a synthetic phantom we have created, inspired by the Shepp-Logan phantom [43], for the purposes of quantitative and qualitative evaluation of segmentation with vicinity prior. Due to the small number of structures and their simplistic spatial arrangement, the Shepp-Logan phantom is not well-suited for such evaluation. The layout of the 10 structures our phantom features (Fig. 3a) mimics the spatial configuration of some thoracic and abdominal organs in a coronal cross-section of a CT scan. In fact, the intensity levels of these structures were derived from mean intensity values of corresponding structures in a CT image.

Graph Cut segmentation with either prior was applied to the pixels<sup>2</sup> as well as to CVT clusterings of  $512 \times 512$  images generated from the phantom in Fig. 3a by the introduction

of additive zero-mean Gaussian noise with standard deviation  $\sigma \in [0, 0.2]$ , image intensity values being clamped to the range  $[0, 1]$ . 40 images were generated for each noise level. Of ranges swept by the hyperparameter  $t$  of (5), we report results in the representative window  $[0, 1]$ . We report clustering resolution as a percentage of the number of image pixels. The set of “seeds” used in likelihood estimation and to impose constraints is given in Fig. 4a. The number of seed-marked pixels for each label is in the range  $[32, 160]$  and roughly proportional to the size of the corresponding structure, otherwise constraints of equal support for all labels would create uneven bias of segmentation ease for smaller structures, such as “9” and “10”.

Fig. 4 gives a qualitative comparison of best results produced by Potts and vicinity prior-based Graph Cut segmentation applied to the pixels and a 10% clustering of an image generated from the phantom with  $\sigma = 0.09$ , which can be considered a relatively high noise level. The vicinity prior was defined according to the graph model presented in Fig. 3b. Improvements brought by the vicinity prior are obvious on both segmentations, especially on thin structures “2” and “9”.

In the remaining of this section as well as in Fig. 4, 5 and 6, any reference to DSM is actually to overall DSM measures for all structures as defined by (9). Mean overall DSM values are calculated on segmentations corresponding to best temperature settings for all 40 images generated at a given noise level.

In Fig. 5a we give mean DSM maps for segmentation with either prior applied to pixels. We calculate mean DSM values for segmentations corresponding to a temperature setting on all images generated at a given noise level. Improvements brought by the vicinity prior are remarkable, particularly for noise levels up to  $\sigma = 0.1$ . Note that low mean DSM values corresponding to low temperature settings are due to high regularization by the prior since data terms would not contribute as much at that level. Also recall that the vicinity prior incurs higher regularization than the Potts prior since its highest penalty in this case is four times greater than its Potts counterpart. To observe the difference between the two priors on a finer scale, we give a magnified view of DSM maps on the low temperature range  $[0.005, 0.05]$ . We do not give mean DSM maps for segmentation applied to clustered images which display similar performance patterns.

Fig. 5b and 6c give mean DSM comparisons of Potts, adjacency and vicinity prior segmentations on pixels and clus-

<sup>2</sup>We have performed some of the evaluations on image pixels (voxels) for comparison purposes with clustered image segmentation. The theoretical framework outlined in Section III still applies by letting variables  $i \in \mathcal{S}$  correspond to image pixels (voxels)  $v \in \mathcal{I}$  and defining a 4-connected (6-connected) neighborhood  $\mathcal{N}_i$  for each variable (respectively in 2D).



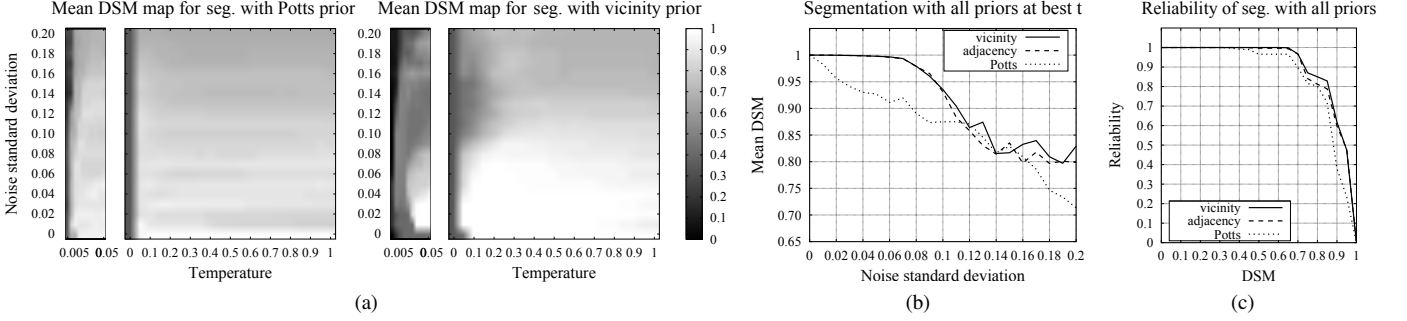


Fig. 5. Quantitative evaluation and comparison of Potts, adjacency and vicinity prior-based pixel segmentation on the synthetic phantom dataset.

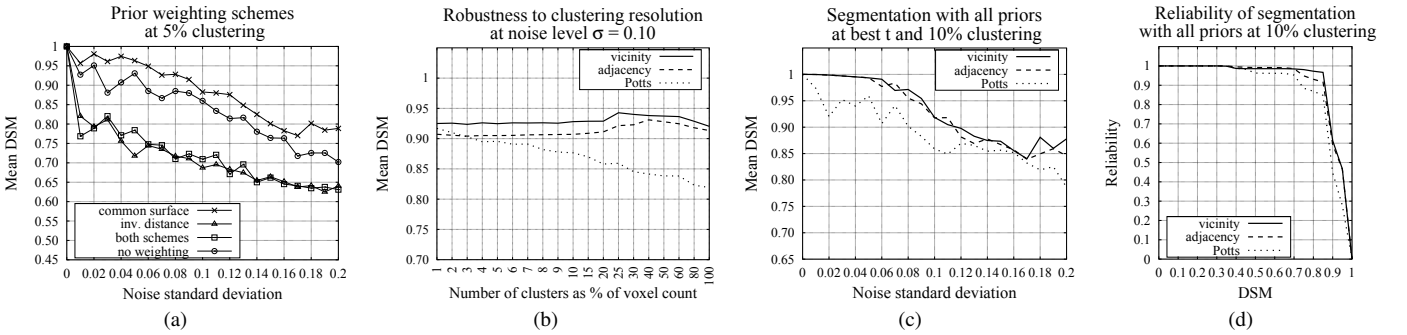


Fig. 6. Quantitative evaluation and comparison of Potts, adjacency and vicinity prior-based clustered image segmentation on the synthetic phantom dataset.

ters with respect to increasing noise levels at the temperature setting yielding the best segmentation for each image. Fig. 5c and 6d compare reliability curves of segmentation with these priors on pixels and clusters respectively. Since evaluations with other clustering resolutions display similar behavior, we present the results of those corresponding to the level 10%.

We also evaluate the robustness of Potts, adjacency and vicinity prior segmentation to clustering resolution. Due to limited space, results corresponding to a single noise level are given in Fig. 6b. We can see that vicinity prior segmentation is quite robust to clustering resolution. This means that similar segmentations can be obtained using lower clustering resolutions thus improving both runtime and memory footprint.

Overall, on evaluations with respect to varying noise and clustering resolution, we notice that the vicinity prior outperforms the Potts prior, and the adjacency prior to a lesser extent.

Lastly, we present the results of evaluating prior weighting schemes for varying clustering resolution. Results corresponding to clustering level 5% are given in Fig. 6a. Let  $C_i$  and  $C_j$  be a pair of clusters sharing a common boundary (a surface in 3D) with centroids  $c_i$  and  $c_j$  respectively. We compare the following four weighting schemes: 1) cluster common-boundary length weighting, where pairwise terms (6) are weighted by  $|\partial C_i \cap \partial C_j|$ , 2) inverse centroid-distance weighting, where pairwise terms are weighted by  $1/d(c_i, c_j)$ , 3) both weighting schemes applied together, and 4) no weighting. Fig. 6a shows that the first weighting scheme is a clear winner.

## B. Simulated and real medical images

1) *BrainWeb simulated MRI dataset*: We have performed a comprehensive quantitative evaluation and comparison on the

BrainWeb simulated MRI dataset [30]. In the creation of this dataset, a semirealistic anatomical head and brain phantom was used by an MRI simulator of T1-, T2- and PD-weighted modalities to generate images according to MR acquisition physics with 6 noise levels (0%, 1%, 3%, 5%, 7% and 9%) and 3 intensity inhomogeneity levels (0%, 20% and 40%). For evaluation purposes, we have used 18 images from the T1-weighted modality covering the entire noise and inhomogeneity ranges. Image volumes are of size  $217 \times 181 \times 217$  voxels with isotropic 1 mm voxel size. The phantom itself used to generate images comprises 10 structure and tissue classes, of which we use 9 classes<sup>3</sup> listed in Table I.

It should be noted that the phantom exhibits certain anatomical inconsistencies. For example, some structures corresponding to tissue classes “muscle” and “skin” are located inside the skull within the volume of tissue class “cerebrospinal fluid” and adjacent to “gray matter” structures. Therefore, we have defined our graph-based anatomical model as a compromise between phantom imprecisions and true anatomical properties.

In Table I, we give means and standard deviations for overall as well as individual-class DSM measures calculated on voxel and 10%-clustered image segmentations corresponding to best temperature settings for all images in the dataset. The vicinity prior is defined according to the graph model in Fig. 8a. It is easy to observe in Table I that Graph Cut segmentation attains better average performance levels with the vicinity prior on the BrainWeb dataset. It also performs slightly better on clusters. Improvements significant at level  $\alpha = 0.05$  are highlighted.

<sup>3</sup>We have merged the tissue class “glial matter” into the class “gray matter” without affecting anatomical consistency.

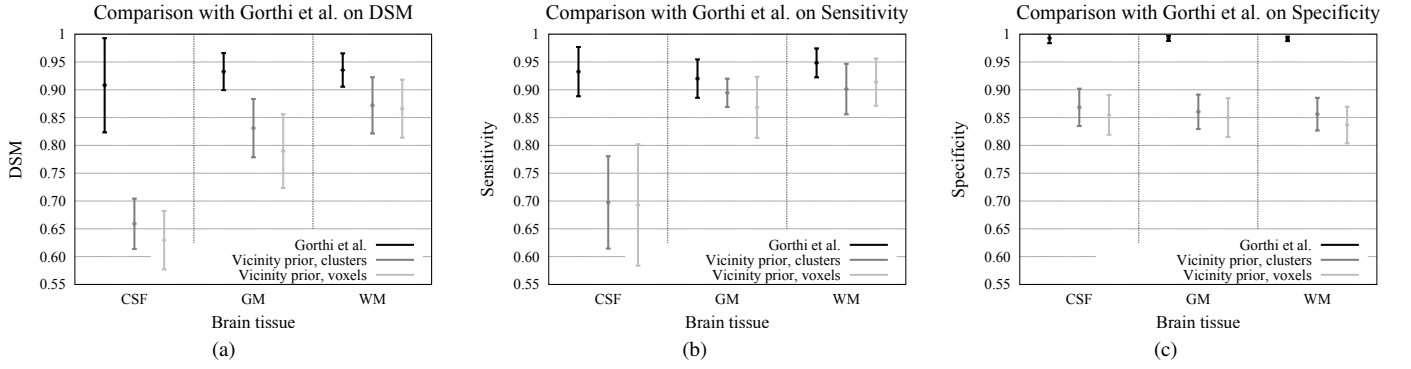


Fig. 7. Quantitative comparison of brain tissue segmentation with [31] on the BrainWeb simulated MRI dataset.

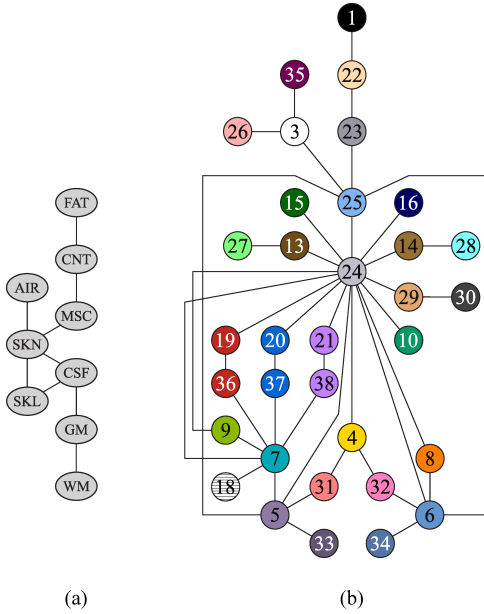


Fig. 8. Graph-based anatomical models of structure adjacency relations for (a) BrainWeb simulated MRI dataset phantom: AIR. background, CNT. connective tissue, CSF. cerebrospinal fluid, FAT. fat, GM. gray matter, MSC. muscle, SKL. skull, SKN. skin, WM. white matter (b) 3D-IRCADb thoracic-abdominal CT image: 1. air 3. bone 4. heart 5. lung (R) 6. lung (L) 7. liver 8. spleen 9. gallbladder 10. pancreas 13. kidney (R) 14. kidney (L) 15. adrenal gland (R) 16. adrenal gland (L) 18. hyperplasia 19. aorta 20. vena cava 21. portal vein 22. skin 23. fat 24. visceral fat 25. muscle 26. marrow 27. renal pyramids (R) 28. renal pyramids (L) 29. digestive tract tissue 30. digestive tract cavity 31. pulmonary arteries (R) 32. pulmonary arteries (L) 33. bronchi (R) 34. bronchi (L) 35. intervertebral discs 36. hepatic arteries 37. hepatic vena cava 38. hepatic portal veins.

Mean runtime and memory footprint figures for voxel and 10%-clustered image segmentations are given in Table III.

We have compared our results on the BrainWeb dataset with those reported in [31] which is the only recent work we are aware of that evaluates the Expansion Moves algorithm on brain tissue classification using the Potts prior model. There are, however, some differences between the segmentation approach in [31] and ours. Firstly, it uses Gaussian likelihood functions the parameters of which are learned during an iterative application of the algorithm. Secondly, it discards all but brain structures i.e. cerebrospinal fluid, gray and white matters, which exhibit good contrast in images in general [44]. Thus,

computational complexity is reduced and the segmentation task is made easier. In our evaluation, we have used the same subset of the BrainWeb dataset as [31] corresponding to the 12 T1-weighted images generated with inhomogeneity levels 20% and 40% and all noise levels. In Fig. 7, we compare means and standard deviations for DSM, sensitivity and specificity measures<sup>4</sup> for the 3 brain structures from full-image vicinity prior segmentations with corresponding measures in [31]. Unsurprisingly, results we report are not superior. Specialized methods are expected to, and indeed should, produce better results. Therefore, the present comparison is an indication of how close a generic approach can come to a specialized one on this problem instance. Nevertheless, a practical advantage of our approach is that no manual tissue removal is required to segment the brain volume.

2) *Thoracic-abdominal CT image*: Lastly, we present a segmentation case study of in vivo acquired data on a thoracic-abdominal 3D CT-scan realized during the arterial phase in inhaled position. The image comes from the 3D-IRCADb dataset [45] and is accompanied by a manual segmentation of 21 structures created by clinical experts of which we use 16 in evaluation. The image volume we use is of size  $480 \times 370 \times 167$  voxels with  $0.961 \times 0.961 \times 1.8$  mm voxel size. We have segmented all visible structures in this image, including those for which no ground-truth estimate was available. The vicinity prior is defined according to the graph model of 34 structures given in Fig. 8b. Seeds were input by a user through a graphical interface by marking few dozens of pixels of any target structure visible in every tenth axial cross-section. Table II gives overall and individual-class DSM measures for Potts and vicinity prior-based Graph Cut segmentation applied to the 5%-clustered image at  $t = 0.5$ . The statistical test shows that the improvement brought by vicinity prior on mean DSM is significant at level  $\alpha = 0.05$ . Qualitative comparison is given in Fig. 9 on coronal cross-sections of ground-truth, vicinity prior and Potts prior-segmented images. Note that the ground-truth can be relied on for qualitative inspection only for the structures listed in Table II. Fig. 10 gives 3D views of simplified surface meshes generated from different structure volumes via the mesh simplification method in [46].

<sup>4</sup>Sensitivity and specificity are defined respectively as the true-positive and the true-negative rates for individual label classes.

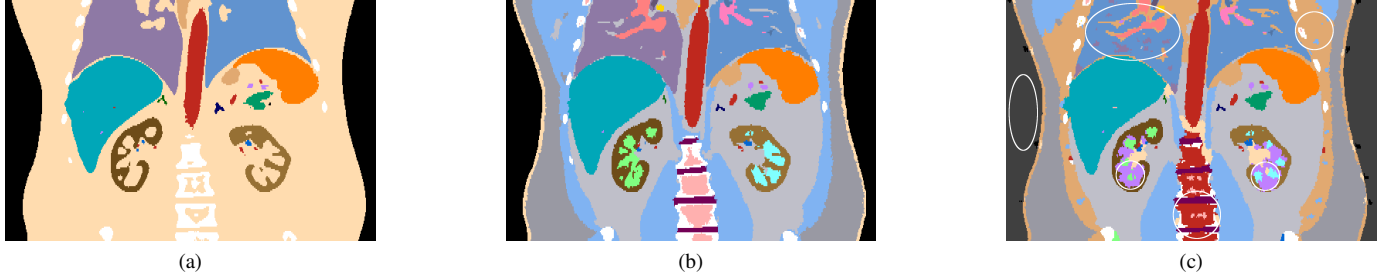


Fig. 9. Qualitative comparison of Potts and vicinity-prior based segmentations of the 5%-clustered 3D-IRCADb thoracic-abdominal CT image. Coronal cross-sections are for (a) ground-truth (b) vicinity prior and (c) Potts prior-segmented images. Refer to color code in Fig. 8b. Ellipses on (c) indicate erroneous segmentations of right lung (top left), muscle (top right), air (center left), right and left renal pyramids (center) and bone marrow (bottom). Notice how right and left lungs are correctly segmented as distinct organs with vicinity prior.

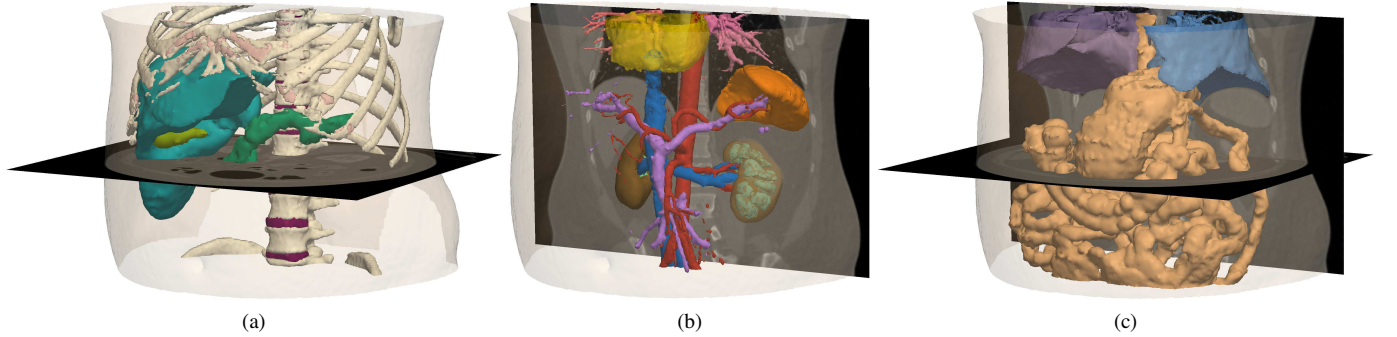


Fig. 10. 3D views of surface meshes for selected structures generated from 3D-IRCADb thoracic-abdominal CT image segmentation via vicinity prior.

Runtime and memory footprint figures for voxel and 5%-clustered image segmentations are given in Table III.

We conclude this section with surface error measurements for Potts and vicinity prior-based segmentations of the 5%-clustered thoracic-abdominal CT image. In Fig. 11, we trace the cumulative histograms of error (11) for 6 representative structures from both segmentations. Presented graphs indicate the percentage of segmented surface lying within a given distance from the true surface. We can see that for reasonably low error levels of 0–1 cm, vicinity prior segmentation has retrieved a considerably larger fraction of structure boundaries than its Potts prior counterpart suffering much lower maximum error. However, it has produced some noticeable outliers, notably for the heart and right lung volumes. Such outliers constitute a small percentage of the structure, and can be corrected via “corrective” seeds as mentioned in Section III-B.

## V. DISCUSSION AND FUTURE WORK

Unlike the Potts prior model, the vicinity prior model we propose is a piecewise-constant model incurring multiple levels of penalization capturing the explicit spatial configuration of objects in multiobject segmentation problems. Compared to shape prior-based segmentation approaches, structural models capturing such spatial configurations are quite robust to shape deformations, because relative positions of objects remain largely stable. Furthermore, such models are relatively easy to define at an arbitrary level of detail out of simple specifications of spatial relationships on pairs of objects, and may also

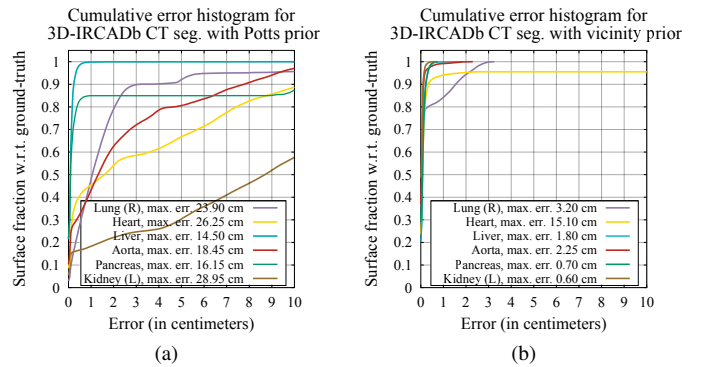


Fig. 11. Surface error measurements for (a) Potts and (b) vicinity prior segmentations of the 5%-clustered 3D-IRCADb thoracic-abdominal CT image.

be generated from existing anatomical models for medical applications as we have seen in Section IV.

The shortest-path formulation produces a prior that is submodular which allows us to optimize (5) via the efficient Expansion Moves Graph Cut algorithm. Pairwise priors learned from training sets might not satisfy this property. Therefore some authors truncate nonsubmodular terms in order to optimize with Graph Cut algorithms [24], and others resort to arbitrary energy minimizers [18].

Qualitative validation and quantitative analysis on synthetic, simulated and real medical images in Section IV confirmed the advantages of our vicinity prior model over the standard Potts model for Graph Cut segmentation, particularly for the

segmentation of distinct objects having identical intensities, the accurate placement of structure boundaries in addition to robustness to clustering resolution. However, like most Graph Cut approaches, our segmentation method is prone to a shrinking bias. This is especially true for thin, elongated structures, like vasculature, where the sum of pairwise costs along the boundary of a vessel is higher than the cost of a shortcut through its thin structure, especially when data-cost terms do not incur a sufficiently strong preference for the vessel's class. We can notice the fragmentary segmentation of abdominal and hepatic arteries and portal veins in Fig. 10b. A number of solutions have been proposed to overcome this problem. For example, [47] improves an initial single-object Graph Cut segmentation with connectivity constraints introduced by the user via markings on image regions which need to be connected to the principal object. [48] proposes a Hessian-based “vesselness” likelihood approach for Graph Cut segmentation of hepatic blood vessels and portal veins in CT images, which could be applied to similar structures such as bronchi and pulmonary arteries.

We are currently investigating the extension of the graph-based prior model to express other types of relations, notably spatial directionality. Since the shortest-path weight formulation on a directed graph would produce asymmetric pairwise prior definitions, such extensions would understandably require revisiting some of the fundamental aspects of the underlying theory of Bayesian labeling.

Another interesting venue for future research is the introduction of uncertainty into the graph-based prior model. As noted in Section III-A, graph models we use to define the vicinity prior are not limited to unit or discrete edge weights. Continuous weights might be used, for example, to express degrees of adjacency, which might be learned from several model variants and thus capture possible variabilities.

TABLE I  
MEAN AND STANDARD DEVIATIONS FOR OVERALL AND INDIVIDUAL-CLASS DSM MEASURES FOR VOXEL AND 10%-CLUSTERED SEGMENTATIONS OF BRAINWEB SIMULATED MRI DATASET WITH POTTS AND VICINITY PRIORS. HIGHLIGHTED VALUES CORRESPOND TO IMPROVEMENTS SIGNIFICANT AT  $\alpha = 0.05$ .

Structures	voxels		clusters 10%	
	Potts	vicinity	Potts	vicinity
Air	0.98 $\pm$ 0.01	0.98 $\pm$ 0.01	0.98 $\pm$ 0.01	0.99 $\pm$ 0.00
Cerebrospinal fluid	0.25 $\pm$ 0.16	<b>0.65</b> $\pm$ 0.06	0.29 $\pm$ 0.11	<b>0.67</b> $\pm$ 0.05
Gray matter	0.79 $\pm$ 0.09	0.82 $\pm$ 0.08	0.80 $\pm$ 0.08	<b>0.85</b> $\pm$ 0.06
White matter	0.86 $\pm$ 0.08	0.88 $\pm$ 0.07	0.87 $\pm$ 0.07	0.88 $\pm$ 0.06
Fat	0.89 $\pm$ 0.05	0.84 $\pm$ 0.08	0.87 $\pm$ 0.04	0.85 $\pm$ 0.06
Muscle	0.73 $\pm$ 0.06	0.74 $\pm$ 0.12	0.73 $\pm$ 0.07	<b>0.79</b> $\pm$ 0.09
Skin	0.70 $\pm$ 0.07	<b>0.75</b> $\pm$ 0.09	0.69 $\pm$ 0.07	<b>0.75</b> $\pm$ 0.07
Skull	0.76 $\pm$ 0.07	0.79 $\pm$ 0.05	0.76 $\pm$ 0.05	<b>0.80</b> $\pm$ 0.04
Connective tissue	0.71 $\pm$ 0.14	0.69 $\pm$ 0.21	0.73 $\pm$ 0.12	0.75 $\pm$ 0.15
Overall	0.81 $\pm$ 0.03	<b>0.85</b> $\pm$ 0.04	0.82 $\pm$ 0.03	<b>0.87</b> $\pm$ 0.04

## REFERENCES

[1] S. Seifert, A. Barbu, S. K. Zhou, D. Liu, J. Feulner, M. Huber, M. Suehling, A. Cavallaro, and D. Comaniciu, “Hierarchical parsing and semantic navigation of full body CT data,” in *SPIE Medical Imaging*, Lake Buena Vista, FL, USA, Feb. 2009.

TABLE II  
OVERALL AND INDIVIDUAL-CLASS DSM MEASURES FOR 5%-CLUSTERED SEGMENTATION OF 3D-IRCADb THORACIC-ABDOMINAL CT IMAGE WITH POTTS AND VICINITY PRIORS. HIGHLIGHTED VALUES CORRESPOND TO IMPROVEMENTS SIGNIFICANT AT  $\alpha = 0.05$ .

Structures	Potts	vicinity	Structures	Potts	vicinity
Bone	0.735	0.857	Kidney (R)	0.907	0.967
Heart	0.918	0.988	Kidney (L)	0.880	0.968
Lung (R)	0.096	0.976	Adrenal gland (R)	0.765	0.743
Lung (L)	0.646	0.980	Adrenal gland (L)	0.706	0.707
Liver	0.990	0.989	Hyperplasia	0.645	0.646
Spleen	0.965	0.982	Aorta	0.512	0.950
Gallbladder	0.902	0.923	Vena Cava	0.894	0.923
Pancreas	0.895	0.959	Portal vein	0.537	0.800
Overall	0.721	0.962	<i>mean</i>	0.750	<b>0.897</b>

TABLE III  
RUNTIME OF VICINITY PRIOR SEGMENTATION ON VOXELS AND 10% AND 5%-CLUSTERED BRAINWEB MR AND 3D-IRCADb CT IMAGES RESPECTIVELY, RECORDED RESPECTIVELY ON A 4-CORE CPU @ 2.84 GHZ WITH 6 GB OF RAM AND A 16-CORE CPU @ 2.66 GHZ WITH 47 GB OF RAM. LAST 4 COLUMNS GIVE CLUSTERING, GRAPH CONSTRUCTION AND OPTIMIZATION TIMES IN MINUTES AND MEMORY FOOTPRINT IN MB.

Image	Segmentation	Clust.	Constr.	Optim.	Mem.
BrainWeb MRI	voxels	—	0.25	12.46	3053.13
	clusters 10%	3.19	0.18	2.83	376.55
3D-IRCADb CT	voxels	—	2.10	460.23	23864.38
	clusters 5%	15.48	1.44	14.57	2624.69

[2] G. Fouquier, J. Atif, and I. Bloch, “Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations,” *Computer Vision and Image Understanding*, vol. 116, no. 1, pp. 146–165, 2012.

[3] T. Okada, Y. Yoshida, M. Hori, R. M. Summers, Y.-W. Chen, N. Tomiyama, and Y. Sato, “Abdominal multi-organ segmentation of CT images based on hierarchical spatial modeling of organ interrelations,” in *Proc. of the 3rd Int. Conf. on Abdominal Imaging: Computational and Clinical Applications*, 2012, p. 173–180.

[4] M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, “Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT,” *Medical Image Analysis*, vol. 16, no. 4, pp. 904–914, 2012.

[5] A. Blume, W. Chun, D. Kogan, V. Kokkevis, N. Weber, R. W. Petterson, and R. Zeiger, “Google body: 3D human anatomy in the browser,” in *ACM SIGGRAPH 2011 Talks*, 2011, p. 19:1–19:1.

[6] C. Rother, V. Kolmogorov, and A. Blake, ““GrabCut”: interactive foreground extraction using iterated graph cuts,” in *ACM SIGGRAPH*, New York, NY, USA, 2004, pp. 309–314.

[7] V. Vezhnevets and V. Konouchine, ““Grow-Cut” - interactive multi-label N-D image segmentation,” in *Graphicon*, 2005, pp. 150–156.

[8] Y. Boykov and G. Funka-Lea, “Graph cuts and efficient N-D image segmentation,” *Int. J. of Comput. Vision*, vol. 70, no. 2, pp. 109–131, 2006.

[9] D. M. Greig, B. T. Porteous, and A. H. Seheult, “Exact maximum a posteriori estimation for binary images,” *J. of the Roy. Stat. Soc. Series B (Methodological)*, vol. 51, no. 2, pp. 271–279, 1989.

[10] H. Ishikawa, “Exact optimization for markov random fields with convex priors,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1333–1336, 2003.

[11] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.

[12] A. Delong and Y. Boykov, “A scalable graph-cut algorithm for N-D grids,” in *26th IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, Jun. 2008.

[13] V. Vineet and P. Narayanan, “CUDA cuts: Fast graph cuts on the GPU,” in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops*, Jun. 2008, pp. 1–8.



- [14] H. Lombaert, Y. Sun, L. Grady, and C. Xu, "A multilevel banded graph cuts method for fast image segmentation," in *10th IEEE Int. Conf. on Computer Vision*, vol. 1, Oct. 2005, pp. 259–265.
- [15] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy snapping," *ACM Trans. on Graphics*, vol. 23, no. 3, p. 303–308, Aug. 2004.
- [16] C. Cigla and A. A. Alatan, "Efficient graph-based image segmentation via speeded-up turbo pixels," in *17th IEEE Int. Conf. on Image Processing*, Hong Kong, Hong Kong, Sep. 2010, pp. 3013–3016.
- [17] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *IEEE 12th Int. Conf. on Computer Vision*, Oct. 2009, pp. 670–677.
- [18] A. Lucchi, Y. Li, X. Boix, K. Smith, and P. Fua, "Are spatial and global constraints really necessary for segmentation?" in *2011 IEEE Int. Conf. on Computer Vision*, Nov. 2011, pp. 9–16.
- [19] Z. Song, N. Tustison, B. Avants, and J. Gee, "Adaptive graph cuts with tissue priors for brain MRI segmentation," in *3rd IEEE Int. Symp. on Biomedical Imaging*, Apr. 2006, pp. 762–765.
- [20] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, Los Alamitos, CA, USA, 2005, pp. 755–762.
- [21] N. Vu and B. Manjunath, "Shape prior segmentation of multiple objects with graph cuts," in *26th IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, Jun. 2008.
- [22] J. Malcolm, Y. Rathi, and A. Tannenbaum, "Graph cut segmentation with nonlinear shape priors," in *IEEE Int. Conf. on Image Processing*, vol. 4, Oct. 2007, pp. 365–368.
- [23] X. Liu, O. Veksler, and J. Samarabandu, "Order-preserving moves for graph-cut-based optimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1182–1196, 2010.
- [24] J. Winn and J. Shotton, "The layout consistent random field for recognizing and segmenting partially occluded objects," in *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, Jun. 2006, pp. 37–44.
- [25] A. Delong and Y. Boykov, "Globally optimal segmentation of multi-region objects," in *12th Int. Conf. on Computer Vision*, Kyoto, Japan, Sep. 2009, pp. 285–292.
- [26] M. Cuadra, L. Cammoun, T. Butz, O. Cuisenaire, and J. Thiran, "Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images," *IEEE Trans. Med. Imag.*, vol. 24, no. 12, pp. 1548–1565, Dec. 2005.
- [27] H. Jacinto, R. Kéchiachian, M. Desvignes, R. Prost, and S. Valette, "A web interface for 3d visualization and interactive segmentation of medical images," in *17th Int. Conf. on 3D Web Technology*, 2012, pp. 51–58.
- [28] A. Levinstein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "TurboPixels: fast superpixels using geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2290–2297, 2009.
- [29] O. Veksler, Y. Boykov, and P. Mehrani, "Superpixels and supervoxels in an energy optimization framework," in *11th European Conf. on Computer Vision*, Heraklion, Crete, Greece, Sep. 2010, pp. 211–224.
- [30] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, 1998.
- [31] S. Gorthi, J. Thiran, and M. B. Cuadra, "Comparison of energy minimization methods for 3-D brain tissue classification," in *18th IEEE Int. Conf. on Image Processing*, Sep. 2011, pp. 57–60.
- [32] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.
- [33] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [34] J. Dardenne, S. Valette, N. Siauve, N. Burais, and R. Prost, "Variational tetrahedral mesh generation from discrete volume data," *The Visual Computer*, vol. 25, no. 5, pp. 401–410, May 2009.
- [35] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. of the 8th Int. Conf. of Computer Vision*, vol. 2, July 2001, pp. 416–423.
- [36] S. Z. Li, *Markov random field modeling in image analysis*. Springer, 2009.
- [37] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, 1984.
- [38] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "TextronBoost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context," *Int. J. of Comput. Vision*, vol. 81, no. 1, pp. 2–23, 2009.
- [39] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for markov random fields with smoothness-based priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1068–1080, 2008.
- [40] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, 2004.
- [41] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [42] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 163–169, Jul. 1987.
- [43] L. Shepp and B. Logan, "The fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. 21, no. 3, pp. 21–43, Jun. 1974.
- [44] S. Gorthi, private communication, Jul. 2012.
- [45] IRCAD France. (2012) 3D-IRCADb. [Online]. Available: <http://www.ircad.fr/software/3Dircadb/3Dircadb.php>
- [46] S. Valette and J. Chassery, "Approximated centroidal voronoi diagrams for uniform polygonal mesh coarsening," *Computer Graphics Forum*, vol. 23, no. 3, pp. 381–389, 2004.
- [47] S. Vicente, V. Kolmogorov, and C. Rother, "Graph cut based image segmentation with connectivity priors," in *26th IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, Jun. 2008.
- [48] V. Pamulapati, B. Wood, and M. Linguraru, "Intra-hepatic vessel segmentation and classification in multi-phase CT using optimized graph cuts," in *IEEE Int. Symp. on Biomedical Imaging*, 2011, pp. 1982–1985.



**Razmig Kéchiachian** received his BSc and MSc degrees in computer science from the University of Aleppo, Syria and the National Institute of Applied Sciences (INSA) of Lyon, France, respectively in 2007 and 2009. He is currently a PhD candidate in the Images and Models project in the Creatis Laboratory (CNRS UMR 5220, Inserm U1044) at INSA-Lyon. His research interests include biomedical image processing, MRF modeling and combinatorial optimization.



**Sébastien Valette** received his MSc degree from the Electrical Engineering Department at the National Institute for Applied Sciences (INSA) of Lyon, France, in 1998. He obtained the PhD Degree at INSA-Lyon in 2002. He is currently a CNRS researcher in the Images and Models project in the Creatis Laboratory (CNRS UMR 5220, Inserm U1044) at INSA-Lyon. His research interests include geometry and image processing.



**Michel Desvignes** received his MSc degree from the Engineering School of Caen (ISMRA) in 1985 and the PhD degree from the University of Caen in 1990. He is currently a computer science professor at Grenoble Institute of Technology (Grenoble-INP). His teaching and research interests include image processing, pattern recognition, shape modeling and machine learning.





**Rémy Prost** (M'82) received his doctorate degree in Electronics Engineering and his "Docteur ès Sciences" degree from Lyon University and the National Institute of Applied Sciences (INSA), Lyon, France, in 1977 and 1987 respectively. He is currently a professor in the Department of Electrical Engineering at INSA-Lyon. Both his teaching and research interests include digital signal processing, inverse problems, image data compression, multiresolution algorithms, wavelets, shape modeling and meshes processing. He leads the Images and Models project in the Creatis Laboratory (CNRS UMR 5220, Inserm U1044) at INSA-Lyon.